**Computer Engineering Dept.**

**Group No :- 22**

**Project Synopsis (2018-19) - Sem VII**

Analysis and Prediction of child mortality in India based on various parameters.

Mrs Pallavi Saindane

Assistant Professor ,CMPN

Rushikesh Gawande Shreiya Indulkar Himani Keswani Malvika Khatri

V.E.S.I.T V.E.S.I.T V.E.S.I.T V.E.S.I.T

[2015rushikesh.gawande@ves.ac.in](mailto:2015rushikesh.gawande@ves.ac.in) [2015shreiya.indulkarlkar@ves.ac.in](mailto:2015neha.prabhavalkar@ves.ac.in) [2015himani.keswani@ves.ac.in](mailto:2015himani.keswani@ves.ac.in) [2015malvika.khatri@ves.ac.in](mailto:2015malvika.khatri@ves.ac.in)

**Abstract**

Extracting patterns in rate of child mortality based on attributes such as: Malaria, Pneumonia, Diarrhoea, etc. Reducing the mortality rate and increasing Health and Mortality concerns is the aim. Data mining is the science of extracting the useful information from a large amount of data sets or databases that leads to statistical and logical analysis and looking for patterns that could help the decision makers.

**Introduction**

For ensuring proper enforcement of the policies, it is necessary to obtain frequent and accurate results of the areas. Traditional methods like census and various surveys usually consumes many days or even months to analyze the areas. Not only time, manpower but also financial resources would be needed in abundance to carry out such surveys swiftly. So a need of an automated system is felt. Using such a system, we can save time as well as money hence pouring those resources for development in that sector.

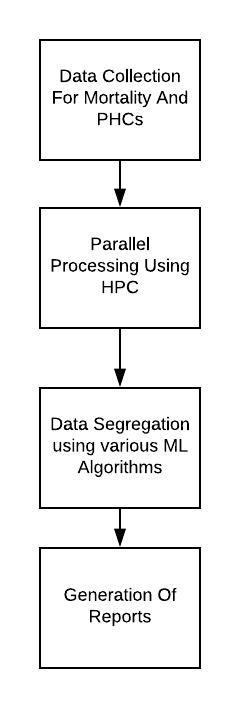
**Problem Statement**

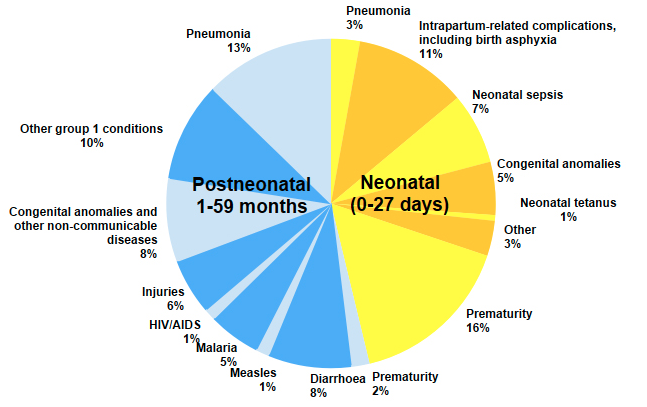
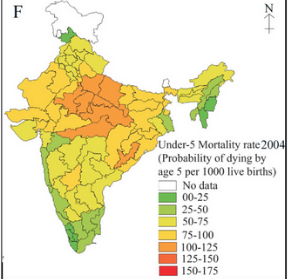
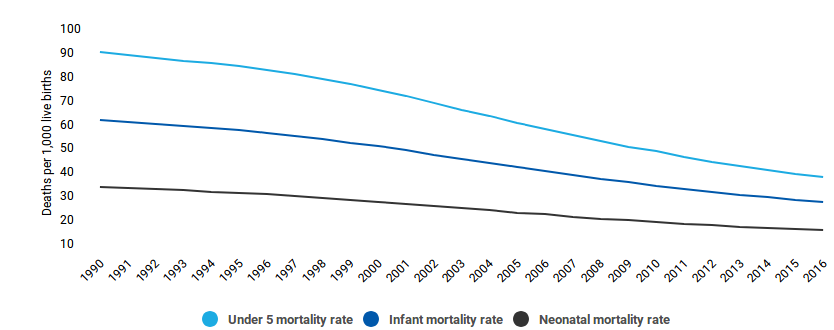
Elimination of child mortality has been the concern since many decades and the greatest global challenge faced by the world and a core requirement for sustainable development, especially for developing countries like India. There are many organizations to help reduce child mortality and many policies implemented to help the nation overcome child mortality, but one of the main problem is finding the impact of various attributes on mortality as well as variation with respect to different areas so major actions can be taken to improve the living environment.

**Solution**

The solution is to perform detailed analysis on child mortality based on various attributes. To develop a predictive model to analyze child mortality rate for a given particular instances. If a child has inappropriate health attributes then there’s a high possibility that he/she could be a case of child mortality.Also,the rate of child mortality and affecting issues can be monitored through analysis. The detailed study of the PHCs per districts available and how those affect the overall mortality of the nation. We will estimate and propose the ideal number of PHCs required for a particular given region.

**Block Diagram:**



**Software and Hardware Requirements:**

Software Requirements: Python based Computer Vision and Machine Learning libraries as well as R programming will be exploited for the development and experimentation of the project. Tools such as Anaconda with Spyder IDE will be utilized for this process.

Hardware Requirements: PC with following specifications:

1. RAM: 8GB or more
2. Hard Disk: 1TB approx
3. Processor: Intel Core i5 or higher
4. Operating System: Ubuntu 18.04

**Methodology**

Step 1: Data collection and data set preparation

Collection of data based on child mortality according to various attributes.

Data segregation using HPC

Step 2: Developing algorithms

Applying ML Algorithms for classifying mortality rate over areas in a period of time and various affecting attributes.

A machine learning model, would be developed for mortality rate prediction based on a single attribute.

Step 3: Training and experimentation on datasets

Training prediction model using ML algorithm.

The model would be trained on training set and the results will be verified.

Step 4: Analysis on real life scenario

Testing model for accuracy

The model would then be tested on test set and the results will be verified using data available for the particular region.

**Experimental Design**

**Dataset**

Dataset will consist of mortality rates based on various attributes and distribution of mortality over different areas.

**Evaluation Measures**

Measuring the accuracy of prediction model.

Measuring deviation of the final output from the actual output.

**Conclusion**

Our analysis is able to anticipate rate of mortality based on various attributes and mortality rate in a particular area using datasets of that area by applying machine learning algorithms on them.

Also,prediction of adequate no. of PHCs in a given area is done by training the district PHCs data. The data generated by the system will be useful to Data Analyst and PHCs in forming the policies. Data can also be used by the government.

**References**

1. “*Determinants of infant and child mortality in Zimbabwe: Results of multivariate hazard analysis”* J Kembo, JK Van Ginneken - Demographic Research, 2009 - JSTOR.
2. *“Neonatal mortality prediction using real-time medical measurements”*Jeff Gilchrist, Colleen M. Ennett, Monique Frize, Erika Bariciak Published in 2011 IEEE International Symposium on Medical Measurements and Applications.